University of Colorado, Boulder
CU Scholar

Computer Science Technical Reports | Computer Science

Spring 5-1-1990

Resource Discovery and Related Research at the University of Colorado ; CU-CS-508-91

Michael F. Schwartz
University of Colorado Boulder

Resource Discovery and Related Research at the University of Colorado[1]

Michael F. Schwartz

CU-CS-508-91

University of Colorado at Boulder
DEPARTMENT OF COMPUTER SCIENCE

[1]This material is based upon work supported in part by NSF cooperative agreement DCR-8420944, and by a grant from AT&T Bell Laboratories.

Abstract

This paper surveys the current state of resource discovery research efforts at the University of Colorado. Work in related areas is also discussed, including distributed collaboration, network management, network integration and network measurement. Pointers to project papers that cover these topics in more depth are included.

Introduction

For the past three years, the Networked Resource Discovery Project at the University of Colorado, Boulder has explored a number of experimental means by which users can discover the existence of resources in a large internet environment. Example resources include network services, documents, retail products, current events, and people. This problem is important because it is an enabling (and currently limiting) technology for the larger issue of supporting *distributed collaboration*, or the accomplishment of tasks through sharing resources among many interrelated individuals across administrative boundaries. Without the ability to discover resources of interest, users perceive only a very limited fraction of the full potential for sharing resources and collaborating with colleagues.

We impose three key goals on our approaches to resource discovery. First, we consider very large environments, spanning national or international networks. Such environments place stringent scalability requirements on the algorithms that can be used. Second, we want to avoid imposing artificial constraints on the resource space organization. Traditional directory services (such as the CCITT X.500 standard [CCITT 1988]) rely on hierarchical organization to achieve good scalability. Unfortunately, the organization of a hierarchy becomes convoluted as an increasingly wide variety of resources is registered, and requires users to understand how the (increasingly deeply) nested components are arranged. Finally, we wish to minimize the need for global administrative agreement over protocols, information formats, and organizational structures. While standards are helpful, it is difficult to specify standards that are both globally adopted and technologically current. Moreover, standards based on a hierarchical organization require a high degree of agreement over the organization of at least the upper levels of the tree. As an increasingly diverse collection of institutions contribute to the global information infrastructure, smooth evolution will require the ability to support multiple organizational structures, and to interoperate with a heterogeneous set of protocols and information formats.

During the course of our research, we have found that the techniques we have developed apply to a number of network problems beyond resource discovery and distributed collaboration, including network management, network integration, and network measurement. We overview a range of our research efforts below.

Internet "White Pages" Tool

The most mature and applied subproject we have completed to date is an Internet "white pages" directory tool called "netfind" [Schwartz & Tsirigotis 1991a]. Given the name of a user and a rough description of where the user works (e.g., the company name or city), netfind attempts to locate telephone and electronic mailbox information about that user. This research focuses on the ability to use a number of existing protocols and highly decentralized sources of relatively unstructured information. Using decentralized information avoids difficult problems of consistency and transfer of authority that are inherent to mechanisms that rely on building auxiliary databases to hold resource information, a point we will return to shortly. The ability to use simply structured information is important in heterogeneous, administratively decentralized environments, where global agreement about highly structured information formats is difficult to achieve.

The technique we have developed to support these characteristics involves building an understanding of the semantics of a particular resource discovery application into the algorithms that support searches. In the current case, the particular application is Internet white pages, and the semantically cognizant mechanism is as follows.

We begin with a database of "seed" data, which provides hints of potential machines to probe when a search is requested. This database is built by gathering information from the headers of USENET [Quarterman & Hoskins 1986] news messages over time. These headers typically list the user name, organization name, city, and electronic mailbox for users who post messages. When a search is requested, the seed database is consulted to locate the names of a number of machines associated with institution keywords specified in the search request. Requests use the format "*UserString InstString [InstString ...]*", where *UserString* identifies the user (typically by last name), and the conjunction of one or more *InstStrings* identify the institution where the user works. For example, a search could be requested for "schwartz university colorado" or "schwartz boulder".

If the machines found in the seed database fall within more than three naming domains (an example of one domain being "colorado.edu"), the user is asked to select at most three domains to search. The Domain Naming System [Mockapetris 1987] is then contacted, to locate authoritative name server hosts for each of these domains. The idea is that these hosts are often central administrative machines, with accounts and/or mail forwarding information for many users at a site. Each of these machines is then queried using the Simple Mail Transfer Protocol (SMTP) [Postel 1982], in an attempt to find mail forwarding information about the specified user. If such information is found, the located machines are then probed using the "finger" protocol [Zimmerman 1990]. The results from finger searches can sometimes yield other machines to search as well. Ten lightweight threads are used to allow DNS, SMTP, and finger lookups to proceed in parallel, to increase resilience to host and network failures.

Because many different institutional keywords will lead to the same seed database records and Domain information, it is usually quite easy to "guess" keywords that will succeed for any particular search. Moreover, netfind can often find a user even if the remote site does not support all of the above protocols, or if some steps in the protocol sequence fail. For example, if finger is disabled because of security concerns, mail forwarding information may sometimes still be found. Or, if no mail forwarding information is found, netfind attempts to finger some of the machines matched from the seed database. This ability to function in the presence of failures or partial remote protocol support is an example of a technique for supporting fault tolerant resource discovery without global agreement. We utilize this technique to a more significant extent in our network visualization project, described later in this paper.

Netfind's tolerance of partial remote protocol support allows it to locate information about a large proportion of Internet users. Measurements indicate that the scope of the directory is upwards of 1,147,000 users in 1,929 sites. This scope is significantly larger than other existing Internet directory services, which require that users register with an administratively centralized service (as with the SRI Network Information Center WHOIS service [Harrenstien, Stahl & Feinler 1985]), or that special servers be run at many sites around the Internet (as with X.500).

Netfind's ability to use highly decentralized information sources allows it to locate very timely information about users. Unlike services that use an auxiliary database that must be updated by a separate administrative procedure, netfind probes the machines on which users do their daily computing. To the best of the author's knowledge, all other white pages services (including WHOIS and X.500) depend on some form of auxiliary database. Populating and keeping such a database up-to-date are difficult tasks.

Using information where it naturally resides is a principle carried forward from our earlier Heterogeneous Name Service work [Schwartz, Zahorjan & Notkin 1987]. However, netfind uses much more decentralized information than the HNS did. The HNS was essentially a framework for supporting users in specifying the semantic operations needed to incorporate new auxiliary database-style name services into a global name service. Moreover, the HNS was used for mapping named objects to data about those objects (such as the network address of a host), rather than for discovering resources. More recently, Droms used an architecture similar to the HNS in his Knowbot Information Service, to provide a white pages service [Droms 1990].

Netfind is in active use by researchers at approximately 50 institutions worldwide, and is being developed further commercially. Measurements of these users indicate that an average search uses approximately 137 packets. While this is more costly than a registration-style directory like X.500, we believe it is quite a reasonable price to pay for providing timely information without global cooperation, particularly when one considers the capacity of next-generation high speed networks.

Electronic Mail Study

Another completed but less "tool oriented" project is a measurement study of global electronic mail communication patterns [Schwartz & Wood 1990]. In this study, we sought to understand how people take part in distributed collaboration, and how distributed collaboration might be better supported. While the main form of distributed collaboration is currently electronic mail, the possibility exists for a much more significant degree of sharing. Electronic mail is primarily used in a point-to-point fashion, supporting the interchange of messages between pairs of individuals. We are interested in a more concurrent, symmetrical style of collaboration involving many participants. For example, a more powerful sharing mechanism could be modeled on the types of interactions that take place at conferences and other meetings, where people discuss issues collectively with people they had not previously met, but who are known to have closely related interests. Electronic "bulletin board" services such as USENET support a crude form of this collaboration, but their restrictive organizational structure (a small number of relatively statically defined interest groups) and means of information distribution (unsequenced, full scale broadcast) do not readily facilitate high quality collaboration.

As a point of departure, we became interested in the organization of human social networks. Such networks use a non-hierarchical organizational structure that scales well. Rather than forming contacts with each other based on a hierarchy, people often establish more direct "networks", by contacting knowledgeable intermediaries who can quickly refer them to other relevant people, cutting across bureaucratic boundaries. For example, by contacting a computer science professor or network manager, someone interested in high-speed networking technology can quickly meet other people who share this interest. These people can, in turn, introduce the person to others who perhaps more closely share his/her particular interests. At the same time, the newcomer can be instrumental in pointing out individuals who share other interests with the people he/she meets.

To study these organizational properties, we collected mail logs from 15 sites around the U.S. and Western Europe for two months. This data constituted a graph containing approximately 50,000 users in 3,700 different sites around the world. Applying graph theoretic analyses yielded a number of insights about how users collaborate by electronic mail. We found that the average path between people is short, which is a statistically rendered version of the small diameter postulated by the so-called "small world" phenomenon [Travers & Milgram 1969]. This property indicates that the graph can support rapid information dissemination. Furthermore, we found that the graph edges are highly redundant, indicating that the graph can support reliable information dissemination. These properties are highly sought in computer networks, yet arise naturally in human social networks.

From the perspective of distributed collaboration, an even more interesting characteristic of human social networks is their flexible organizational structure. Rather than forcing all internode relationships to conform to a hierarchy, the graph structure allows individuals to be related to one another through multiple groupings that we call *Specialization Subgraphs (SSGs).* An SSG is a subset of nodes that share common attributes, and that has a small diameter. As an example, in a graph of relationships among people, one SSG could connect individuals based on a shared interest in a particular computer science speciality, a second SSG could connect individuals based on shared responsibilities at a place of employment, and a third SSG could group individuals based on shared cultural/recreational interests. Any individual can belong to many different SSGs, and can search for information about a particular topic by consulting the appropriate SSG.

To measure properties of SSGs, we developed an algorithm to cluster individuals by shared interests without access to the contents of the mail messages, by computing properties of the mail interconnection graph using traffic analysis techniques [Callimahos 1989]. We found it necessary to apply the traffic analysis to a subgraph derived by a graph reduction technique that eliminated "noise" caused by our statistical sampling. The combination of these two techniques allowed us to derive lists of individuals who share interests with one or more specified individuals. (We ran the computation specifying a number of individuals whose interests were known to us to validate the procedure.) This clustering algorithm has potential applications for distributed collaboration (discussed further in the Section "Internet Resource Mapping/Discovery Project" below), as well as privacy implications for electronic mail. Moreover, applying this algorithm to each of a randomly chosen set of 500 nodes within the graph provided measurements of how people collaborate, indicating the existence of a large number of different but heavily interrelated groupings of individuals based on shared interest, and underscoring the importance of supporting a number of different organizational structures for distributed collaboration.

Probabilistic "Yellow Pages" Protocols

Another project on which we have made substantial progress is a study of probabilistic algorithms that construct and search a resource graph that supports attribute-based ("yellow pages") specifications [Schwartz 1989, Schwartz 1990]. For this project we assume it is desirable to find a small number of instances of a moderately large class of objects. For example, in searching for a supplier of a particular piece of computer hardware, finding 5 out of 100 suppliers in a metropolitan area would often suffice. We also assume that it is acceptable to return different answers to the same query across search sessions. If consistent responses to queries are desired, one could build a front-end user interface that cached results, and provided identical responses across search sessions.

Based on these assumptions, we designed and simulated a protocol to support a set of *agents* in organizing and searching the resource space. Agents maintain pointers to sources of resource information, and access these sources via intermediary *brokers*, which enforce the access control policies and encapsulate the heterogeneity of the information repositories. While agents are intended to be part of the network infrastructure, each broker belongs to the organization whose resource information it exports.

While brokers are an important part of the model, the main focus of our research is on the agent protocols for organizing and searching the resource space. Rather than having an administrative body specify how the space is organized, agents organize the space dynamically, according to what resources exist and the types of searches that users make. Agents use a probabilistic *Sparse Diffusion Multicast* primitive to disseminate information about resources at uniformly distributed, randomly chosen nodes around the network, and likewise to route search requests randomly around the network. Sparse Diffusion Multicast can be implemented with simple modifications to Cheriton and Deering's Internet multicast protocol [Cheriton & Deering 1990].

Randomly disseminating resource information is intended to place the information within a reasonably small neighborhood of any agent in the network, so that during searches it is likely that the information can be found using simple random probes. Since the types of resources that exist and the searches users request are not random, a cache management policy is used to prefer graph edges between agents that maintain related information, to form specialization subgraphs. Using this policy, a search initiated at a random agent may cause some random search behavior at the start of the search, until a member of an appropriate specialization subgraph is reached. If such a subgraph is reached, searches proceed in a more directed fashion. If a user continues to use the same agent, over time that agent will maintain pointers to sources of the type of information for which that user often searches.

Our results to date indicate that this approach can support a non-hierarchical resource space for an environment roughly the size of a country, with several thousand sites participating in resource registration and searches. We are currently extending these results to model various aspects of the protocol in more detail.

The probabilistic nature of this approach also supports fair access among competing information providers, an important issue in commercial environments such as the U.S. telecommunications industry [Greene 1988] or computer reservation systems [Gifford & Spector 1984].

Internet Resource Mapping/Discovery Project

A project currently under way is an effort to support resource discovery in decentralized environments of such scale that the resource space cannot be completely organized. In such an environment, mechanisms are needed to support incremental organization of the resources, based on the efforts of many geographically distributed individuals, and a range of different information sources of varying degrees of quality. Our approach to this problem is to use mechanisms that "tap into" existing network protocols and information sources to provide an immediately useful tool (much as netfind did), supplemented by mechanisms that allow users to superimpose additional organization on the resource space in an incremental fashion. As a concrete test case, we are developing a prototype that focuses on public Internet archive sites accessible via the "anonymous" File Transfer Protocol [Postel & Reynolds 1985]. This is an interesting test case, because it encompasses thousands of administratively decentralized sites containing a collection of resources of considerable practical value.

The part of this effort associated with tapping into existing infrastructure is now basically complete [Schwartz et al. 1990]. In this prototype, three levels of information quality are supported. At the highest level, resources are described using archive-site-resident databases, with individual resources described according to their conceptual roles. Below that, per-user and per-user-site caches are maintained, to record resources that have been found by individual users during their explorations. At the lowest level, the system scans USENET electronic bulletin board articles using a simple set of heuristics to recognize announcements about public archive sites, to provide a simple keyword-based index of resources throughout the Internet.

To support users in superimposing additional organization on the resource space, we are currently developing a system architecture that allows any individual or group of users who share common interests to build a structure (called a *view*) that superimposes organization on the resource space according to their particular interests. For example, a group interested in graphics might build a view that organizes the world according to PostScript, Tools, Window Systems, Images, and Discussions, with pointers to network accessible resources of these various types nested into this structure. A view is intended to be a simple structuring mechanism for loosely integrating an administratively decentralized pool of resources, with properties much like those of specialization subgraphs. Views can include pointers to parts of other views, so that related interest groups (such as people interested in operating systems and people interested in data communications) can cross reference each others' views. Views are not constrained to be hierarchical, although in practice many links will probably be tree-like.

Because an arbitrary number of different views may coexist, our research investigates how to support users in discovering, constructing, and sharing views. We consider several problems. Views must be replicated and distributed, to improve availability and performance. Views must be secured against accidental or malicious modification. It must be easy to update and reorganize views. Finally, it must be easy to search for resources.

We place particular focus on supporting searches. Our approach is to build a simple flat index of each view. While flat searching has not worked well in some situations involving large scale (such as library information systems), searching a view in this manner should work well because any particular view will be fairly small and highly focused in scope. Because of this, it should be relatively easy to "guess" appropriate keywords for searching a view, without the difficulty of keywords matching many unrelated subjects. Moreover, the underlying space in a view is structured. Hence, a user could examine a subtree in a view once a match occurs, unlike the corresponding situation with the flat underlying spaces supported by information retrieval systems.

Of course, the problem remains that a user must discover appropriate views to search when trying to find a resource. To address this problem, an area of future research we intend to pursue involves experimenting with a means of automatically interelating views, based on the interest clustering algorithm developed in connection with our electronic mail study. Doing so will provide a dynamically evolving set of links between related views, and allow users to search for resources without having to know what views exist.

We will deploy the prototype at a number of Internet sites, and enlist the participation of users in building a distributed map of the resources available via anonymous FTP. Sites interested in participating in this effort when the software becomes stable are encouraged to contact the author.

Network Visualization Project

Another project currently under way involves using aggressive resource discovery techniques to support visualization of characteristics of the global Internet, such as topology, congestion, routing, and protocol usage [Schwartz et al. 1991]. As with netfind and the anonymous FTP prototype, we will use a number of protocols and information sources, to support discovery in the absence of global agreement on any one protocol or information source. For the visualization project, however, we will use a much more extensive collection of protocols and information sources, including the Address Resolution Protocol [Plummer 1982], the Internet Control Message Protocol [Postel 1981], the Domain Naming System [Mockapetris 1987], the Simple Network Management Protocol [Case et al. 1989], and a dozen others. Essentially, this project is exploring the extent to which one can integrate a heterogeneous, administratively decentralized network by using aggressive resource discovery techniques.

We will deploy a collection of servers around the Internet, each of which periodically executes a set of discovery protocols to maintain information about its local internet. An X-window user interface will allow users to connect to servers and browse the state of the Internet. Servers will cache information and support queries that allow partial retrieval of remote information, for example retrieving information that has changed since the last update to the local cache contents. Servers will also be parameterized so that system administrators may chose what discovery protocols will execute and their scheduling frequency, according to their perceptions of the relative importance of network loading, timliness, and completeness of the data.

Wide Area Demand Resource Distribution Project

We have recently begun work on a mechanism for widely distributing files without broadcasting them to all nodes on the network (as is the case with the Network News Transfer Protocol [Kantor & Lapsley 1986]), and without causing files to be retrieved multiple times across individual network links (as is the case with anonymous FTP). The basic idea, suggested by Phil Karn of Bell Communications Research, is to distribute files in response to requests for them, caching them at intermediate nodes along a dynamically-developed spanning tree of the Internet. Doing so can potentially reduce the load on the Internet substantially, since FTP currently accounts for 45% of the bytes transmitted on the NSFNET backbone [NSF Network Service Center 1989]. Moreover, this mechanism could underly a network-transparent mode of resource sharing, whereby resources are named independently of the particular hosts that hold them. This technique will be best supported by an Internet multicast mechanism [Cheriton & Deering 1990], but will also function without such specialized support.

A number of issues must be addressed in this project, including accountability of cache contents (to ensure that intermediary sites do not accidentally or maliciously modify file replicas), cache consistency (to ensure that the most recent version of a distributed file is retrieved), and the proper level of file fragmentation, in response to alternative routing of file contents across the dynamically routed Internet.

Summary

This paper presents a brief overview of a number of projects exploring techniques to support resource discovery and distributed collaboration. We focus on techniques appropriate for national or international networks, and the concomitant issues of scalability, administrative decentralization, and supporting natural organization. Our efforts involve wide area distributed prototypes and measurement studies, using the global TCP/IP Internet as an experimental testbed. In addition to resource discovery and distributed collaboration, the techniques we have developed also apply to a number of related problems, including network integration, network management, and, more generally, to supporting globally distributed applications [Schwartz & Tsirigotis 1991b].

Supporting resource discovery raises some difficult issues concerning privacy of information. While security mechanisms may be imposed to preserve privacy in some cases, in many cases such mechanisms are either difficult to provide, or of questionable merit. We believe that privacy is essentially a social issue, and as such requires careful consideration about the policies that will manage the technical solutions, rather than an emphasis simply on technical solutions to security problems. Moreover, we believe that the best way to understand the tension between privacy and resource discovery is to explore the issues raised by building and deploying resource discovery prototypes.

Interested readers may obtain copies of many of the project related papers by anonymous FTP from latour.colorado.edu, in the directory pub/RD.Papers, or by contacting the author by electronic mail at schwartz@latour.colorado.edu.

Bibliography

[CCITT 1988]
CCITT. The Directory, Part 1: Overview of Concepts, Models and Services. ISO DIS 9594-1, CCITT, Gloucester, England, December 1988. Draft Recommendation X.500.

[Callimahos 1989]
L. D. Callimahos. *Traffic Analysis and the Zendian Problem.* Aegean Park Press, Laguna Hills, California, 1989.

[Case et al. 1989]
J. Case, M. Fedor, M. Schoffstall and C. Davin. A Simple Network Management Protocol (SNMP). Request For Comments 1098, April 1989.

[Cheriton & Deering 1990]
D. R. Cheriton and S. E. Deering. Multicast Routing in Datagram Internetworks and Extended LANs. *ACM Transactions on Computer Systems*, 8(2), pp. 85-110, May 1990.

[Droms 1990]
R. E. Droms. Access to Heterogeneous Directory Services. Proceedings of the InfoCom Conference, June 1990.

[Gifford & Spector 1984]
D. Gifford and A. Spector. The TWA Reservation System. *Communications of*

the ACM, 27(7), pp. 650-665, July 1984.

[Greene 1988]
H. H. Greene. United States of America, Plaintiff, v. Western Electric Company, Inc., et al., Defendants. Civil Action No. 82-0192, U.S. District Court, District of Columbia, March 1988. Triennial review of Modified Final Judgement divesting AT&T of the Regional Bell Operating Companies.

[Harrenstien, Stahl & Feinler 1985]
K. Harrenstien, M. Stahl and E. Feinler. NICName/Whois. Request For Comments 954, October 1985.

[Kantor & Lapsley 1986]
B. Kantor and P. Lapsley. Network News Transfer Protocol - A Proposed Standard for the Stream-Based Transmission of News. Request For Comments 977, U.C. San Diego and U.C. Berkeley, February 1986.

[Mockapetris 1987]
P. Mockapetris. Domain Names - Concepts and Facilities. Request For Comments 1034, USC Information Sciences Institute, November 1987.

[NSF Network Service Center 1989]
NSF Network Service Center. *Internet Resources Guide*. NSF Network Service Center, August 1989. Available by anonymous FTP from nnsc@nnsc.nsf.net.

[Plummer 1982]
D. C. Plummer. An Ethernet Address Resolution Protocol -- Or -- Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware. Request For Comments 826, November 1982.

[Postel 1981]
J. Postel. Internet Control Message Protocol. Request For Comments 792, USC Information Sciences Institute, September 1981.

[Postel 1982]
J. B. Postel. Simple Mail Transfer Protocol. Request For Comments 821, USC Information Sciences Institute, August 1982.

[Postel & Reynolds 1985]
J. Postel and J. Reynolds. File Transfer Protocol (FTP). Request For Comments 959, USC Information Sciences Institute, October 1985.

[Quarterman & Hoskins 1986]
J. S. Quarterman and J. C. Hoskins. Notable Computer Networks. *Communications of the ACM*, 23(10), pp. 932-971, October 1986.

[Schwartz, Zahorjan & Notkin 1987]
M. F. Schwartz, J. Zahorjan and D. Notkin. A Name Service for Evolving, Heterogeneous Systems. *Proceedings of the Eleventh ACM Symposium on Operating Systems Principles*, pp. 52-62, Austin, Texas, November 1987. Published as Operating Systems Review 21(5).

[Schwartz 1988]
M. F. Schwartz. Autonomy vs. Interdependence in the Networked Resource Discovery Project. Position paper, ACM SIGOPS European Workshop, Cambridge, England, September 1988.

[Schwartz 1989]
M. F. Schwartz. The Networked Resource Discovery Project. *Proceedings of the IFIP XI World Congress*, pp. 827-832, San Francisco, California, August 1989.

[Schwartz et al. 1990]
M. F. Schwartz, D. R. Hardy, W. K. Heinzman and G. Hirschowitz. Supporting Resource Discovery Among Public Internet Archives Using a Spectrum of Information Quality. Technical Report CU-CS-487-90, September 1990. Submitted for publication.

[Schwartz & Wood 1990]
M. F. Schwartz and D. C. M. Wood. A Measurement Study of Organizational Properties in the Global Electronic Mail Community. Technical Report CU-CS-482-90, Department of Computer Science, University of Colorado, Boulder, Colorado, August 1990. Submitted for publication.

[Schwartz 1990]
M. F. Schwartz. A Scalable, Non-Hierarchical Resource Discovery Mechanism Based on Probabilistic Protocols. Technical Report CU-CS-474-90, Department of Computer Science, University of Colorado, Boulder, Colorado, June 1990. Submitted for publication.

[Schwartz et al. 1991]
M. F. Schwartz, D. H. Goldstein, R. K. Neves and D. C. M. Wood. An Architecture for Discovering and Visualizing Characteristics of the Global Internet. Technical Report, Department of Computer Science, University of Colorado, Boulder, Colorado, 1991. In preparation.

[Schwartz & Tsirigotis 1991a]
M. F. Schwartz and P. G. Tsirigotis. Experience with a Semantically Cognizant Internet White Pages Directory Tool. To appear, *Journal of Internetworking Research and Experience*, 1991.

[Schwartz & Tsirigotis 1991b]
M. F. Schwartz and P. G. Tsirigotis. Principles for Supporting Globally Distributed Applications. Technical Report, Department of Computer Science, University of Colorado, Boulder, Colorado, 1991. In preparation.

[Travers & Milgram 1969]
J. Travers and S. Milgram. An Experimental Study of the Small World Problem. *Sociomety*, 32(4), pp. 425-443, 1969.

[Zimmerman 1990]
D. Zimmerman. The Finger User Information Protocol. Request For Comments 1194, Center for Discrete Mathematics and Theoretical Computer Science, November 1990.